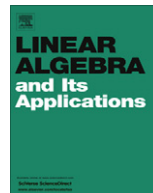




ELSEVIER

Contents lists available at SciVerse ScienceDirect

Linear Algebra and its Applications

journal homepage: www.elsevier.com/locate/laa

Patterns of alternating sign matrices

Richard A. Brualdi^{a,*}, Kathleen P. Kiernan^a, Seth A. Meyer^a,
Michael W. Schroeder^b^a Department of Mathematics, University of Wisconsin, Madison, WI 53706, United States^b Department of Mathematics, Marshall University, Huntington, WV 25755, United States

ARTICLE INFO

Article history:

Received 14 April 2011

Accepted 1 March 2012

Available online xxxx

Submitted by N. Shaked-Monderer

In admiration, to Avi Berman, Moshe
Goldberg, and Raphi Loewy

AMS classification:

05B20

05C22

05C50

15B36

Keywords:

(Symmetric) alternating sign matrix

ASM

Pattern

Row (column) sum vector

Alternating signed graph

Term rank

ABSTRACT

We initiate a study of the zero–nonzero patterns of $n \times n$ alternating sign matrices. We characterize the row (column) sum vectors of these patterns and determine their minimum term rank. In the case of connected alternating sign matrices, we find the minimum number of nonzero entries and characterize the case of equality. We also study symmetric alternating sign matrices, in particular, those with only zeros on the main diagonal. These give rise to alternating signed graphs without loops, and we determine the maximum number of edges in such graphs. We also consider $n \times n$ alternating sign matrices whose patterns are maximal within the class of all $n \times n$ alternating sign matrices.

© 2012 Elsevier Inc. All rights reserved.

1. Introduction

An *alternating sign matrix*, henceforth abbreviated ASM, is an $n \times n$ $(0, +1, -1)$ -matrix such that the $+1$ s and -1 s alternate in each row and column, beginning and ending with a $+1$. For example,

* Corresponding author.

E-mail addresses: brualdi@math.wisc.edu (R.A. Brualdi), kiernan@math.wisc.edu (K.P. Kiernan), smeyer@math.wisc.edu (S.A. Meyer), schroederm@marshall.edu (M.W. Schroeder).

$$\begin{bmatrix} 0 & 0 & +1 & 0 & 0 \\ 0 & +1 & -1 & +1 & 0 \\ +1 & -1 & +1 & -1 & +1 \\ 0 & +1 & -1 & +1 & 0 \\ 0 & 0 & +1 & 0 & 0 \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} 0 & 0 & 0 & +1 & 0 \\ +1 & 0 & 0 & -1 & +1 \\ 0 & +1 & 0 & 0 & 0 \\ 0 & 0 & 0 & +1 & 0 \\ 0 & 0 & +1 & 0 & 0 \end{bmatrix}$$

are ASMs. Since every permutation matrix is an ASM, ASMs can be thought of as generalizations of permutation matrices. The following elementary properties of ASMs follow from their definition:

- (i) The number of nonzeros in each row and column is odd.
- (ii) The first and last nonzero in each row and column is a +1.
- (iii) The first and last rows and columns contain exactly one +1 and no -1s.
- (iv) The partial row sums of the entries in each row, starting from the first (or last) entry, equal 0 or 1. A similar property hold for columns. Moreover, all row and column sums equal 1.

Reversing the order of the rows or columns of an ASM results in another ASM. More generally, thinking of an $n \times n$ matrix as an $n \times n$ square, applying the action of an element of the dihedral group of order 8 to an ASM results in another ASM. A matrix obtained from an ASM by row and column permutations is generally not an ASM, even in the case of a simultaneous row and column permutation. In particular, being an ASM is not a combinatorial property of a matrix.

The substantial interest in ASMs in the mathematics community originated from the alternating sign matrix conjecture of Mills et al. [4] in 1983; see [1,5,6] for a history of this conjecture and its relation with other combinatorial constructs. This conjecture, proved by Zeilberger [7] and also by Kuperberg [3], asserts that the number of $n \times n$ ASMs equals

$$\frac{1!4!7! \cdots (3n - 2)!}{n!(n + 1)!(n + 2)! \cdots (2n - 1)!}$$

In this paper we are concerned with the zero–nonzero pattern of an ASM.

An $n \times n$ ASM A has a unique decomposition of the form

$$A = A_1 - A_2,$$

where A_1 and A_2 are $(0, +1)$ -matrices without any +1 in the same position. For example,

$$\begin{bmatrix} 0 & +1 & 0 & 0 \\ +1 & -1 & +1 & 0 \\ 0 & 0 & 0 & +1 \\ 0 & +1 & 0 & 0 \end{bmatrix} = \begin{bmatrix} 0 & +1 & 0 & 0 \\ +1 & 0 & +1 & 0 \\ 0 & 0 & 0 & +1 \\ 0 & +1 & 0 & 0 \end{bmatrix} - \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & +1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}. \tag{1}$$

The *pattern* of an ASM A is the $(0, +1)$ -matrix $\tilde{A} = A_1 + A_2$. The ASM in (1) has pattern

$$\begin{bmatrix} 0 & +1 & 0 & 0 \\ +1 & +1 & +1 & 0 \\ 0 & 0 & 0 & +1 \\ 0 & +1 & 0 & 0 \end{bmatrix},$$

which, dropping the symbol 1 or dropping both symbols 1 and 0, we also write as

$$\begin{bmatrix} 0 & + & 0 & 0 \\ + & + & + & 0 \\ 0 & 0 & 0 & + \\ 0 & + & 0 & 0 \end{bmatrix} \quad \text{or} \quad \begin{bmatrix} & + & & \\ + & + & + & \\ & & & + \\ & + & & \end{bmatrix}.$$

An $n \times n$ ASM A can be viewed as a signed biadjacency matrix of a bipartite graph $G \subseteq K_{n,n}$, equivalently, the biadjacency matrix of a signed bipartite graph, where the vertices in each part of the bipartition are linearly ordered and where the edges are signed ± 1 . Similarly, an $n \times n$ symmetric ASM can be viewed as the signed adjacency matrix of a loopy¹ graph $G \subseteq K_n$, equivalently, the adjacency matrix of a signed loopy graph, where the vertices are linearly ordered and where the edges are labeled ± 1 . As appropriate, we use the terminology *alternating signed bipartite graph*, abbreviated ASBG, and *alternating signed loopy graph*, abbreviated ASLG. If there are only 0s on the main diagonal of a symmetric ASM, that is, there are no loops in the graph, then we have an *alternating signed graph*, abbreviated ASG.

Our main goal in this paper is to investigate possible patterns of ASMs and to determine their properties. In Section 2 we discuss some basic properties of patterns of ASMs. The maximum number of nonzeros in an $n \times n$ ASM (an ASBG) is easily determined along with the case of equality. The minimum number of nonzero entries in an $n \times n$ ASM is n , since permutation matrices are ASMs. Let A be an $n \times n$ ASM, and let $R = (r_1, r_2, \dots, r_n)$ and $S = (s_1, s_2, \dots, s_n)$ be the row sum vector and column sum vector of the pattern of A . Then the components of R and S are odd positive integers with $r_1 = s_1 = r_n = s_n = 1$ and

$$r_1 + r_2 + \dots + r_n = s_1 + s_2 + \dots + s_n.$$

We characterize the possible row sum (column sum) vectors of patterns of ASMs. In Section 3 we determine the minimum number of nonzeros in an ASM whose associated signed bipartite graph is connected. Recall that the *term rank* $\rho(X)$ of a matrix X is the maximum possible number of nonzeros in X with no two from the same row and column, and that by the König–Egérvary theorem, $\rho(X)$ is the minimum number of rows and columns of X that contain all the nonzeros of X . The term rank of an $n \times n$ ASM may be n , as the permutation matrices show. In Section 4 we determine the smallest possible term rank of an $n \times n$ ASM. In Section 5 we determine the maximum number of edges in an ASG of order n , that is, the maximum number of nonzeros in an $n \times n$ symmetric ASM with only 0s on the main diagonal. We also determine when equality occurs. In Section 6 we consider ASMs which are maximal in the sense that it is not possible to change one or more 0 to ± 1 resulting in another ASM.

Some concluding remarks are given in Section 7.

2. Patterns of ASMs

We first observe that there is a simple construction which shows that given any $k \times l$ $(0, +1, -1)$ -matrix B , there is an ASM that contains B as a submatrix. To construct such an ASM, we proceed as follows:

- If some row of B is a zero row, we put in a new column anywhere which has a $+1$ in that row and 0s elsewhere.
- If in B the first nonzero entry of some row is a -1 , we put in a new column anywhere to its left whose only nonzero entry is a $+1$ in that row.
- If the last nonzero entry of some row of B is a -1 , we put in a new column anywhere to its right whose only nonzero entry is a $+1$ in that row.

¹ Loopy refers to the fact that loops are allowed but not required.

- (d) If B has two -1 s in row i , with no nonzeros in between, we put in a new column anywhere in between the two -1 s whose only nonzero entry is a $+1$ in that row i .
- (e) If B has two $+1$ s in row i with no nonzeros in between, we put in a new column anywhere between the two $+1$ s and two new rows, one anywhere above row i and one anywhere below row i . The new column has a -1 in row i and a $+1$ in the two new rows.
- (f) We repeat (a)–(e) using the columns of B in place of the rows.

Applying this construction, we obtain an ASM A containing B as a submatrix; we say that such an A is obtained from B by an *elementary ASM expansion* of B . In general, an ASM containing B as a submatrix is an *ASM expansion* of B . It follows from our construction that there can be no forbidden submatrix condition that can be used to characterize ASMs.

Now let $A = [a_{ij}]$ be an $n \times n$ ASM with pattern \tilde{A} . Let $R = (r_1, r_2, \dots, r_n)$ and $S = (s_1, s_2, \dots, s_n)$ be the row sum vector and column sum vector of \tilde{A} .

Let \mathbf{j}_n equal the n -vector $(1, 1, \dots, 1)$ of all 1s, and let \mathbf{k}_n be the n -vector $(1, 3, 5, \dots, 5, 3, 1)$ which reads backward the same as it reads forward. Thus $\mathbf{k}_6 = (1, 3, 5, 5, 3, 1)$ and $\mathbf{k}_7 = (1, 3, 5, 7, 5, 3, 1)$. There is a well-known special $n \times n$ ASM D_n which we call the *diamond ASM* and illustrate in (2) for $n = 6$ and 7 , it being obvious how to generalize to arbitrary n :

$$D_6 = \begin{bmatrix} & & & & & \\ & & & & + & \\ & & & & + & - & + \\ & & + & - & + & - & + \\ + & - & + & - & + & \\ + & - & + & & & \\ & & + & & & \\ & & & & & \end{bmatrix}, \quad D_7 = \begin{bmatrix} & & & & & & + \\ & & & & & & + & - & + \\ & & & & + & - & + & - & + \\ + & - & + & - & + & - & + & \\ + & - & + & - & + & & \\ & & + & - & + & & \\ & & & & + & & \\ & & & & & & \end{bmatrix}. \tag{2}$$

For n even, the matrix obtained from D_n by taking its rows in the reverse order is also an ASM and we also refer to it as a diamond ASM. Note that a diamond ASM of order n is symmetric, and its row and column sum vectors of D_n equal \mathbf{k}_n .

Let $n = 4k$ be a multiple of 4. For later use we introduce another $n \times n$ symmetric ASM E_{4k} of order $4k$ which we call a *near-diamond ASM*. The pattern matrix \widetilde{E}_{4k} has row and column sum vector

$$(1, 3, 5, 7, \dots, 4k - 5, 4k - 3, 4k - 3, 4k - 3, 4k - 3, 4k - 5, \dots, 7, 5, 3, 1),$$

which differs from the row and column sum vector of \widetilde{D}_{4k} only in the middle two entries where \widetilde{D}_{4k} has $4k - 1$. Thus E_{4k} has four fewer non zeros than D_{4k} . The unique $+$ in row 1 of E_{4k} is in column $2k + 2$ (where the unique $+$ in row 1 of D_{4k} is in column $2k + 1$). We illustrate the construction of E_{4k} for $k = 2$ and compare it with the construction of D_8 in (3) below.

$$E_8 = \begin{bmatrix} & & & & & & + \\ & & & & + & - & + \\ & & & + & - & + & - & + \\ & & + & - & + & - & + \\ + & - & + & - & + & & \\ + & - & + & & & & \\ & & + & & & & \\ & & & & & & \end{bmatrix} \quad \text{and} \quad D_8 = \begin{bmatrix} & & & & & & + \\ & & & & + & - & + \\ & & + & - & + & - & + \\ + & - & + & - & + & - & + \\ + & - & + & - & + & & \\ & & + & - & + & & \\ & & & & + & & \\ & & & & & & \end{bmatrix}. \tag{3}$$

Lemma 2.1. *If A is an $n \times n$ ASM whose pattern \tilde{A} has row sum vector R and column sum vector S , then $\mathbf{j}_n \leq R, S \leq \mathbf{k}_n$ (entrywise).*

Moreover, $R = S = \mathbf{j}_n$ if and only if A is a permutation matrix, and $R = S = \mathbf{k}_n$ if and only if A is a diamond ASM.

Proof. The assertions involving \mathbf{j}_n are obvious since an ASM must have at least one nonzero in each row and column, and permutation matrices are ASMs. The assertions involving \mathbf{k}_n are also straightforward. The first and last row (respectively, column) of an ASM can contain only one $+1$ and no -1 s. Let $i > 1$. The -1 s in row i of an ASM can only be in those columns for which the corresponding column sum of the leading $(i - 1) \times n$ submatrix A_{i-1} of A equals 1. Since the row sums of A_{i-1} equal 1 and the column sums of A_{i-1} are nonnegative, exactly $i - 1$ of the column sums of A_{i-1} are equal to 1. We conclude that there are at most $i + 1$ s and at most $(i - 1) - 1$ s in row i , and hence at most $2i - 1$ nonzeros in row i . The same argument applies to the column sums taken from bottom to top. This proves that $R \leq \mathbf{k}_n$, and in a similar way we get that $S \leq \mathbf{k}_n$.

If n is odd and $R = S = \mathbf{k}_n$, then row $(n + 1)/2$ has only nonzero entries, alternating between $+1$ and -1 , and the ASM is uniquely determined as D_n . If n is even, then rows $(n/2) \pm 1$ contain exactly one 0. Because $R = \mathbf{k}_n$, this unique 0 must be in the first column of one of these rows and the last column of the other, and we again get that A is D_n . \square

The linear constraints $\mathbf{j}_n \leq R, S \leq \mathbf{k}_n, r_1 = r_n = s_1 = s_n = 1$, and $r_1 + r_2 + \dots + r_n = s_1 + s_2 + \dots + s_n$ for vectors $R = (r_1, r_2, \dots, r_n)$ and $S = (s_1, s_2, \dots, s_n)$ of positive odd integers are not sufficient to guarantee the existence of an $n \times n$ ASM whose pattern has row sum vector R and column sum vector S . For example, if $n = 5$ and $R = S = (1, 1, 5, 1, 1)$, there is no 5×5 ASM A whose pattern has this row and column sum vector. This is because the third row of such an A must be $[+ \ - \ + \ - \ +]$ and the third column must be the transpose of this vector. Hence the row sum and column sum vector of such an A must be \mathbf{k}_5 and not $(1, 1, 5, 1, 1)$. This leads to another necessary condition for integral n -vectors $R = (r_1, r_2, \dots, r_n)$ and $S = (s_1, s_2, \dots, s_n)$, namely,

$$\frac{r_i - 1}{2} \leq |\{j : s_j \geq 3\}| \quad (1 \leq i \leq n).$$

This is because row i of an ASM contains $(r_i - 1)/2 - 1$ s and so there are at least this many columns with at least 3 nonzeros. This kind of argument can be carried further. Consider consecutive rows i and $i + 1$ of an ASM. Their -1 s occur in different columns. Thus there are at least

$$\frac{r_i - 1}{2} + \frac{r_{i+1} - 1}{2}$$

columns with a -1 and thus at least this many columns with at least three nonzeros. Therefore

$$\frac{r_i - 1}{2} + \frac{r_{i+1} - 1}{2} \leq |\{j : s_j \geq 3\}| \quad (1 \leq i \leq n).$$

There do not seem to be any easily formulated necessary and sufficient conditions for R and S to be the row and column sum vectors of an ASM. However, the inequalities $\mathbf{j}_n \leq R \leq \mathbf{k}_n$ for an n -vector R of positive odd integers are sufficient to guarantee the existence of an $n \times n$ ASM whose pattern has row sum vector R .

Theorem 2.2. *Let $R = (r_1, r_2, \dots, r_n)$ be a vector of positive odd integers. Then there is an $n \times n$ ASM whose pattern has row sum vector R if and only if*

$$\mathbf{j}_n \leq R \leq \mathbf{k}_n. \tag{4}$$

Proof. We note that (4) implies that $r_1 = r_n = 1$. By Lemma 2.1, (4) is satisfied by the row sum vector of the pattern of an $n \times n$ ASM. Now suppose that (4) holds. We show how to construct an ASM

whose pattern has row sum vector R . The ASM is obtained by combining the results of two smaller constructions which we now describe. We first assume that n is an odd integer, $n = 2m + 1$.

Construction, Part I: For $1 \leq k \leq m + 1$ we inductively construct a $k \times (2k - 1)$ $(0, +1, -1)$ -matrix A_k such that the rows satisfy the alternating ± 1 condition of ASMs with the row sums of its pattern equal to (r_1, r_2, \dots, r_k) , and the columns also satisfy the alternating ± 1 condition from top to bottom except for the fact that the full column sum vector is $(1, 0, 1, 0, \dots, 1, 0, 1)$ (and not $(1, 1, \dots, 1)$ as it would be for an ASM). If $k = 1$, then $A_1 = [+]$. If $k = 2$, then

$$A_2 = \begin{bmatrix} 0 & 0 & + \\ + & 0 & 0 \end{bmatrix} \text{ if } r_2 = 1 \quad \text{and} \quad A_2 = \begin{bmatrix} 0 & + & 0 \\ + & - & + \end{bmatrix} \text{ if } r_2 = 3.$$

Suppose that $k \geq 2$ and we have constructed A_k to satisfy our requirements. If $r_{k+1} = 2k + 1$, then we border A_k with a zero column on the left and right, and adjoin the $(2k + 1)$ -vector $[+ - + - \dots + - +]$ as a new row. The resulting $(k + 1) \times (2k + 1)$ -matrix A_{k+1} satisfies our requirements. Now suppose that $r_{k+1} < 2k + 1$. Then we adjoin two zero columns to A_k , a zero column on the left and a zero column between the r_{k+1} and $r_{k+1} + 1$ columns of A_k ; we then adjoin as a new row the $(2k + 1)$ -vector $[+ - + - \dots + - + 0 \dots 0]$. The resulting $(k + 1) \times (2k + 1)$ matrix A_k satisfies our requirements.

In this part of the construction, we finish with an $(m + 1) \times (2m + 1)$ $(0, +1, -1)$ -matrix A_{m+1} satisfying our requirements where the row sum vector of the pattern of A_{m+1} is $(r_1, r_2, \dots, r_{m+1})$. Since the column sum vector of A_{m+1} is $(1, 0, 1, 0, \dots, 1, 0, 1)$, the last nonzero entry, if there is a nonzero entry, in its odd numbered columns is a 1, and in the even numbered columns it is a -1 .

Construction, Part II: Using Construction I we can obtain an $m \times (2m - 1)$ $(0, +1, -1)$ -matrix A'_m , such that the rows of A'_m satisfy the alternating ± 1 condition of ASMs with the row sums of its pattern equal to $(r_{2m+1}, r_{2m}, \dots, r_{m+2})$, and the columns also satisfy the alternating ± 1 condition except for the fact that the full column sums are $(1, 0, 1, 0, \dots, 1, 0, 1)$. Let A''_m be the matrix obtained by bordering A'_m by a zero column on the left and right, and then taking its rows in the reverse order. The column sums of A''_m are $(0, 1, 0, 1, \dots, 0, 1, 0)$. The first nonzero entry of the odd numbered columns of A''_m , if there is one, is a -1 while the first nonzero entry in the even numbered columns of A''_m is a $+1$. It follows that the matrix

$$\begin{bmatrix} A_{m+1} \\ A''_m \end{bmatrix}$$

is a $(2m + 1) \times (2m + 1)$ ASM whose pattern has row sum vector R .

Now assume that n is an even integer $n = 2m$. Then using the ideas above, we obtain an $m \times (2m - 1)$ $(0, +1, -1)$ -matrix whose rows satisfy the alternating ± 1 condition of ASMs with the row sums of its pattern equal to (r_1, r_2, \dots, r_m) and whose columns satisfy the alternating ± 1 condition except for the fact that the full column sums equal $(1, 0, 1, 0, \dots, 1)$. We then append a column of all 0s to obtain an $m \times 2m$ matrix A_m with full column sums equal to $(1, 0, 1, 0, \dots, 1, 0)$. We also obtain an $m \times 2m$ $(0, +1, -1)$ -matrix A''_m whose rows satisfy the alternating ± 1 condition of ASMs with the row sums of its pattern equal to $(r_{2m}, r_{2m-1}, \dots, r_{m+1})$ and whose columns satisfy the alternating ± 1 condition except for the fact that the full column sums equal $(0, 1, 0, 1, \dots, 0, 1)$. Taking the rows of A''_m in the reverse order to produce the matrix A'_m , we obtain the required $2m \times 2m$ ASM

$$\begin{bmatrix} A_m \\ A'_m \end{bmatrix}. \quad \square$$

3. Minimal connected ASMs

Let $A = [a_{ij}]$ be an $n \times n$ ASM, and let $BG(A) \subseteq K_{n,n}$ be the alternating signed bipartite graph (ASBG) determined by A . If $BG(A)$ is connected, then we call A a *connected ASM*; otherwise, A is a *disconnected ASM*. It follows easily that there exist permutation matrices P and Q such that

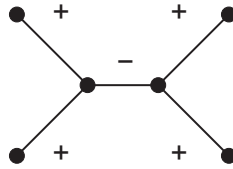


Fig. 1. $BG(D_3)$.

$$PAQ = A_1 \oplus A_2 \oplus \dots \oplus A_h \quad (h \geq 1), \tag{5}$$

where A_1, A_2, \dots, A_h are also ASMs and $BG(A_i)$ is connected for all i . If, for instance, A is a permutation matrix, then in (5) we get $h = n$ and $PAQ = I_n$. In this section we show that the minimum number of nonzero entries of an $n \times n$ connected ASM equals $2n - 1$ (thus $BG(A)$ is a tree) if n is odd, and equals $2n$ (thus $BG(A)$ is a unicyclic graph whose unique cycle has even length) if n is even. In each case, we give a recursive construction to obtain all graphs attaining equality. If A is an $n \times n$ ASM, then the degrees of all the vertices of $BG(A)$ are odd; if $BG(A)$ is a tree, then the sum of the degrees of the vertices in one part of the bipartition equals $2n - 1$, the number of edges of $BG(A)$. Hence $BG(A)$ can be a tree only if n is odd.

Recall that the diamond ASM D_3 equals

$$\begin{bmatrix} & + & & \\ - & & & \\ + & - & + & \\ & & & \\ & & + & \end{bmatrix}.$$

The signed bipartite graph $BG(D_3)$ is shown in Fig. 1.

Let A be an $n \times n$ ASM. Then we use the notation $A * D_3$ to denote an $(n + 2) \times (n + 2)$ matrix obtained from A by identifying a single $+$ of A with a single $+$ of D_3 and inserting two new rows and two new columns in order to embed D_3 in the resulting matrix. Neither the two new rows nor the two new columns need be consecutive, but they must retain the same relative order as in D_3 and must maintain the alternating sign property. The matrix $A * D_3$ is also an ASM, and we say that $A * D_3$ results from A by attaching D_3 . The bipartite graph $BG(A * D_3)$ is obtained by identifying a positive edge of $BG(D_3)$ with a positive edge of $BG(A)$. For instance, if

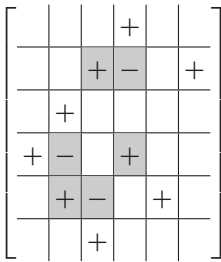
$$A = \begin{bmatrix} & & & + & & \\ & + & - & & + & \\ + & - & + & & & \\ & + & - & + & & \\ & & & & & \\ & & & + & & \end{bmatrix},$$

then one possible $A * D_3$ is the matrix

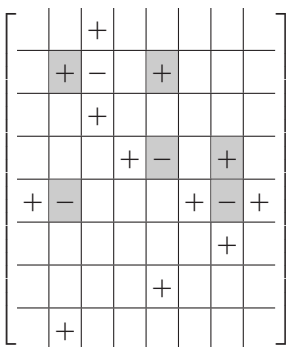
$$\begin{bmatrix} & & & + & & & \\ & + & - & & + & & \\ & & & & & & \\ + & - & + & - & & + & \\ & + & - & & + & & \\ & & & & & & \\ & & & + & & & \\ & & & & & & \\ & & & + & & & \end{bmatrix},$$

where the $+$ of A in its position $(3, 3)$ is identified with the $+$ of D_3 in its position $(2, 1)$, and the two new rows are rows 3 and 6 and the two new columns are columns 4 and 7. Deleting rows 3 and 6 and columns 4 and 7 of $A * D_3$ results in the original ASM A .

Since $BG(D_3)$ is a tree, then by starting with I_1 and attaching D_3 s we can construct for every odd integer n an $n \times n$ connected ASM A with $BG(A)$ a tree. Now suppose that n is an even integer with $n \geq 4$. Then there are $n \times n$ connected ASMs with exactly $2n$ nonzero entries. One such basic family of ASMs is obtained as follows: Let $m \geq 2$ be an integer, and consider a cycle C of even length $2m$ whose edges are (arbitrarily) either positive or negative. Depending on the signs of the edges at each of the individual (original) vertices of C , we attach a new vertex to the graph as well as new edges and vertices as follows: (i) if two positive edges meet at a vertex of C we attach a negative edge to a new vertex and to that new vertex two positive edges attached to two new vertices, (ii) if one positive and one negative edge meet at a vertex of C we attach a positive edge to a new vertex, (iii) if two negative edges meet at a vertex we attach three positive edges to three new vertices. Since the number of edges attached to each vertex of C is odd and C has an even number of vertices, the total number of vertices of the resulting graph is even. If, for instance, the edges of C alternate in sign, then we attach a positive (pendent) edge to each vertex of C . It is easy to see that the resulting graph $G \subseteq K_{2m, 2m}$ can always be realized (in many ways) as a $2m \times 2m$ ASM with exactly $2 \cdot 2m$ nonzero entries; in fact, this construction is just a graphical description of the elementary ASM expansion construction discussed at the beginning of Section 2 (reordering the vertices as necessary). For example, with $m = 3$ and a cycle of length 6 whose edges alternate in sign, we have



With $m = 3$ and a cycle of length 6 with edges $+, +, -, +, -, -$, we have the 8×8 ASM



We call an ASM constructed in this way a *basic unicyclic ASM*.

Given an $n \times n$ connected ASM A with $2n$ nonzero entries, then each matrix of the form $A * D_3$ is an $(n + 2) \times (n + 2)$ connected ASM with $2(n + 2)$ nonzero entries.

Theorem 3.1. *The minimum number of nonzero entries in an $n \times n$ connected ASM equals $2n - 1$ if n is odd and $2n$ if n is even. If n is odd, then an $n \times n$ connected ASM A has exactly $2n - 1$ nonzero entries if and only if A equals the 1×1 identity matrix I_1 or A can be obtained from I_1 by recursively attaching D_3 s.*

If n is even, then an $n \times n$ connected ASM A has exactly $2n$ nonzero entries if and only if A is either a basic unicyclic ASM or can be obtained from a basic unicyclic ASM by recursively attaching D_3 s.

Proof. It follows from our preceding discussion that the minima are as given in the theorem, and the ASMs as described in the theorem attain the minima. We need to show that an ASM satisfying one of the minima can be obtained as described in the theorem.

First consider the case of n odd. We may assume that $n \geq 3$. Let A be an $n \times n$ ASM where $BG(A)$ is a tree T . Let u be any pendent vertex of T and root T at u ; the unique edge at u is positive. Let d be the largest distance of a vertex to u . Since $n \geq 3$, we may assume that $d \geq 3$. The vertices of T are partitioned into sets V_0, V_1, \dots, V_d where V_i consists of all vertices at distance i to u ($0 \leq i \leq d$), and the only edges of T join a vertex in V_i to a vertex in V_{i+1} for some i with $0 \leq i \leq d - 1$. For each i with $1 \leq i \leq d$ and each $v \in V_i$, there is a unique vertex $w \in V_{i-1}$ which is joined by an edge to v .

Let p be any vertex in V_d . Then p is a pendent vertex which is joined by a (positive) edge to a unique vertex $q \in V_{d-1}$. The vertex q must have degree 3, and hence there exist vertices $r \in V_{d-2}$ and $s \in V_d$ with a negative edge joining q to r and a positive edge joining q to s . The vertex r must be incident with at least two positive edges, and so there exists another vertex $t \in V_{d-1}$ such that there is a positive edge joining r and t . Since this edge is positive and $t \in V_{d-1}$, t must be a pendent vertex. It follows that if A' is the $(n - 2) \times (n - 2)$ submatrix of A obtained by deleting the rows or columns corresponding to vertices p, q, s, t , then A is obtained from A' by attaching a D_3 . We need to know that we can choose t and q so that A' is also an ASM, that is, so that the row or column corresponding to vertex r is alternating in sign; there may be many positive and negative edges at vertex r . Since there is only one edge going from r to a vertex in V_{d-3} (one sign in the row or column corresponding to r), there must be such a consecutive pair.

Now assume that $n \geq 4$ is even, and that A is an $n \times n$ connected ASM with $2n$ nonzero entries. Since $BG(A)$ is a connected bipartite graph with $2n$ vertices and $2n$ edges, $BG(A)$ contains a unique cycle and it has even length. The graph $BG(A)$ consists of this cycle with a tree (possibly empty) rooted at each of its vertices. Since A is an ASM, it follows easily that A contains a basic unicyclic ASM U as a submatrix. If $A = U$ we are done. Otherwise, there exists a tree rooted at at least one of the vertices of $BG(U)$. Now, using an argument similar to that used in the odd case, we complete the proof. \square

4. Term rank of ASMs

The term rank $\rho(B)$ of a matrix B is the maximum number of nonzeros of B with no two from the same row or column. The term rank of B also equals the smallest number of rows and columns that contain all the nonzeros of B (see e.g. [2]). An ASM A and its pattern \tilde{A} have the same term rank: $\rho(A) = \rho(\tilde{A})$. As already remarked, the term rank of an $n \times n$ ASM may equal n . In this section we obtain a lower bound for the term rank of an ASM and then provide a construction to show that it is the best possible.

Theorem 4.1. Let A be an $n \times n$ ASM. Then

$$\rho(A) \geq \left\lceil 2\sqrt{n+1} - 2 \right\rceil. \tag{6}$$

Proof. Let the term rank of A be t . Then there exist e rows and f columns of A with $e + f = t$ that contain all the ± 1 s of A . We permute the rows and columns of A bringing these e rows to the top and bringing these f columns to the left, otherwise respecting the order of the rows and columns. The resulting matrix A' has the form

$$\begin{bmatrix} A_1 & A_2 \\ A_3 & O_{n-e,n-f} \end{bmatrix},$$

where A_1 is an $e \times f$ matrix. The columns of A_2 and the rows of A_3 satisfy the alternating sign property of ASMs. Thus the sum of the entries of A_2 equals $n - f$ and the sum of the entries of A_3 equals $n - e$.

Let p_i and n_i denote, respectively, the number of +1s and number of -1s in A_i for $i = 1, 2, 3$. Then, using the alternating sign property, we see that the following relations hold:

$$p_2 - n_2 = n - f, \quad p_3 - n_3 = n - e, \quad (p_1 + p_2) - (n_1 + n_2) = e, \quad (p_2 + p_3) - (n_1 + n_3) = f. \quad (7)$$

The relations in (7) imply that

$$n_1 - p_1 = n - (e + f) = n - t, \quad (8)$$

that is, A_1 has $n - t$ more -1s than +1s. Since A_1 is an $e \times f$ matrix and since $e + f = t$, we obtain

$$n_1 - p_1 = n - t \leq ef \leq \left(\frac{t}{2}\right)^2. \quad (9)$$

Manipulating (9), we get

$$\begin{aligned} n - t &\leq \frac{t^2}{4} \\ 4n + 4 &\leq t^2 + 4t + 4 = (t + 2)^2 \\ 2\sqrt{n + 1} &\leq t + 2 \\ 2\sqrt{n + 1} - 2 &\leq t, \end{aligned}$$

and (6) follows. \square

We now show by construction that for each $n \geq 1$ there exists an $n \times n$ ASM for which equality holds in (6). We consider the two cases:

$$t = 2k \quad \text{if } k(k + 1) < n + 1 \leq (k + 1)^2 \text{ for some integer } k \quad (10)$$

and

$$t = 2k + 1 \quad \text{if } (k + 1)^2 < n + 1 \leq (k + 1)(k + 2) \text{ for some integer } k. \quad (11)$$

First consider the case given by (10), and let B be any $k \times k$ $(0, -1)$ -matrix with exactly $q = (n - 2k)$ -1s. Suppose that for $i = 1, 2, \dots, k$, row i of B has r_i -1s and column i of B has s_i -1s. Then $\sum_{i=1}^k r_i = \sum_{i=1}^k s_i = q$. Applying an elementary ASM expansion to B we obtain an ASM A all of whose nonzero entries are contained in the k rows and k columns that contain its submatrix B . Hence $\rho(A) \leq 2k$. Moreover, the number of rows of A that are not rows of B equals

$$\sum_{i=1}^k (s_i + 1) = q + k = (n - 2k) + k = n - k,$$

and hence A has n rows. Similarly, A has n columns. Since by (10), $2k$ is a lower bound for the term rank of an ASM of order n , we have $\rho(A) = 2k$ as desired.

Now consider the case given by (11), and let B be any $k \times (k + 1)$ $(0, -1)$ -matrix with $q = n - (2k + 1)$ -1s. Again, applying an elementary ASM expansion to B we obtain again an ASM A of order n all of whose nonzero entries are contained in the k rows and $k + 1$ columns of its submatrix B . Hence $\rho(A) \leq 2k + 1$. Since by (11), $2k + 1$ is a lower bound for the term rank, we have $\rho(A) = 2k + 1$ as desired.

This completes the construction that shows that equality can be obtained in Theorem 4.1 for all n . As an example of this construction, let $k = 3$ and $n = 14$ so that we are in the case given by (10). Let B be the 3×3 $(0, -1)$ -matrix with $q = n - 2k = 14 - 6 = 8$ -1s whose only zero is in its lower right corner. Then a matrix produced by an elementary expansion is

(12)

The nonzeros can be covered with the 3 rows and 3 columns containing the -1 s.

Corollary 4.2. *The minimum term rank of an $n \times n$ ASM equals*

$$\lceil 2\sqrt{n+1} - 2 \rceil.$$

5. Symmetric ASMs

The $n \times n$ diamond ASM D_n is a symmetric matrix. As noted in Section 1, we can view an $n \times n$ symmetric ASM $A = [a_{ij}]$ as the adjacency matrix of a signed loopy graph (an ASLG) $G(A)$ with linearly ordered vertices $1, 2, \dots, n$ with a positive (respectively, negative) edge between vertices i and j if and only if $a_{ij} = +1$ (respectively, $a_{ij} = -1$). We denote a signed edge $\{i, j\}$ by $\{i, j\}^+$ and $\{i, j\}^-$, respectively. The defining property of an ASLG is: If i is any vertex and $\{j_1, j_2, \dots, j_k\}$ are the vertices joined to i by an edge (k is necessarily odd) where $1 \leq j_1 < j_2 < \dots < j_k \leq n$, then we have $\{i, j_1, \dots, j_1\}^+, \{i, j_3, \dots, j_3\}^+, \dots, \{i, j_k, \dots, j_k\}^+$ and $\{i, j_2, \dots, j_2\}^-, \{i, j_4, \dots, j_4\}^-, \dots, \{i, j_{k-1}, \dots, j_{k-1}\}^-$. We observe that if A is an $n \times n$ ASM with only zeros on the main diagonal, then n must be even. This is because the degrees of the corresponding loop-free graph $G(A)$ are all odd, and so $G(A)$ has an even number of vertices. Another example of an ASLG with loops is given by the ASM

Its signed loopy graph contains a 5-cycle with vertices 2, 3, 4, 5, 6 and signed edges $\{2, 4\}^-$, $\{4, 6\}^-$, $\{6, 3\}^+$, $\{3, 5\}^-$, $\{5, 2\}^+$, $\{3, 4\}^-$, $\{4, 1\}^+$, and $\{4, 7\}^+$, and positive loops at each of the vertices 2, 3, 4, 5, and 6.

Otherwise, by symmetry, the submatrix of A determined by rows and columns $i - 1, i, i + 1$ is of the form

$$\begin{bmatrix} 0 & e & f \\ e & 0 & g \\ f & g & 0 \end{bmatrix},$$

where $e, f, g \neq 0$. Thus $f = -e$ and $g = -e$, and so $f = g$, violating the alternating sign property of A .

Consider the partition of A given by

$$A = \begin{bmatrix} A_1 & X & Y \\ X^t & A_2 & Z \\ Y^t & Z^t & A_3 \end{bmatrix}, \tag{15}$$

where A_1 and A_3 are $k \times k$ symmetric matrices with zero diagonals, and A_2 is a $2k \times 2k$ symmetric matrix with a zero diagonal. Then, it follows from (14) that A_2 contains at least $k - 1$ zeros above its main diagonal, that is,

$$\sigma^*(A_2) \leq 2k^2 - 2k + 1. \tag{16}$$

For the hollowed-diamond ASM D_{4k}^* , the submatrices corresponding to $A_1, A_3,$ and Y are zero matrices, and the number of zeros above the main diagonal in its $2k \times 2k$ submatrix corresponding to A_2 is exactly k . It follows from the definition of the near-diamond ASM E_{4k} that for the hollowed-near-diamond ASM E_{4k}^* , we have that A_1 and A_3 are zero matrices, the matrix Y has one nonzero entry, a $+$ in its lower left corner, and the number of zeros above the main diagonal in A_2 is $k - 1$.

Now assume that A has the maximum number of nonzeros among all symmetric ASMs of order n with a zero main diagonal. Then since $\sigma(D_{4k}^*) = \sigma(E_{4k}^*) = 8k^2 - 4k$ and $\sigma^*(D_{4k}^*) = \sigma^*(E_{4k}^*) = 4k^2 - 2k$, we have that

$$\sigma(A) \geq 8k^2 - 4k \quad \text{and} \quad \sigma^*(A) \geq 4k^2 - 2k. \tag{17}$$

The number of nonzeros of A in the first k rows (respectively, columns) is at most $1 + 3 + \dots + (2k - 1) = k^2$, and hence, using (16) and (17), we get

$$\begin{aligned} 4k^2 - 2k \leq \sigma^*(A) &\leq \sigma^*(A_2) + k^2 + k^2 - \sigma(Y) \\ &\leq (2k^2 - 2k + 1) + 2k^2 - \sigma(Y) \\ &= 4k^2 - 2k + 1 - \sigma(Y). \end{aligned}$$

From this calculation we see that

$$0 \leq \sigma(Y) \leq 1, \tag{18}$$

and then that the number of nonzeros of A in the first k rows (respectively, columns) is at least $k^2 - 1$, that is,

$$\sigma \left(\begin{bmatrix} A_1 & X & Y \end{bmatrix} \right) \geq k^2 - 1 \quad \text{and, similarly,} \quad \sigma \left(\begin{bmatrix} Y \\ Z \\ A_3 \end{bmatrix} \right) \geq k^2 - 1. \tag{19}$$

Since the number of nonzeros in each row is odd, it follows from (19) that the numbers of nonzeros in rows $1, 2, \dots, k$ of A are $1, 3, \dots, 2k - 1$, respectively, and, similarly, the numbers of nonzeros in the last k columns of A are $2k - 1, \dots, 3, 1$, respectively. This implies that the first k rows of A , after deletion of zero columns, is a $k \times (2k - 1)$ matrix of the form (shown here for $k = 5$)

$$\begin{bmatrix} & & & & + & & & & \\ - & & & & & & & & \\ & & & + & - & + & & & \\ - & & & & & & & & \\ & & + & - & + & - & + & & \\ - & & & & & & & & \\ & + & - & + & - & + & - & + & \\ - & + & - & + & - & + & - & + & \end{bmatrix}. \tag{20}$$

If in A_1 there were a nonzero (necessarily not on its main diagonal), and so a $+1$, then by symmetry there would be an i with $1 \leq i \leq k$, such that there is at least one $+1$ below the main diagonal in row i of A_1 and at least one $+1$ above the main diagonal in column i of A_1 . But then, using the above form for the first k rows of A , there would be a nonzero in the diagonal position (i, i) in A_1 , a contradiction. Thus A_1 , and similarly A_3 , is a zero matrix.

By (18), we have only two cases to consider.

Case 1: $\sigma(Y) = 0$: In this case, since A_1, A_3, Y, Y^t are zero matrices and all row and column sums of A equal 1, the sum of all the entries of A outside of A_2 equals $4k$, and hence the sum of the entries of A_2 is zero. Since A_2 has a zero main diagonal, $\sigma^*(A_2)$ is even. Thus equality cannot occur in (16) and hence $\sigma^*(A_2) \leq 2k^2 - 2k$. Therefore,

$$\sigma^*(A) = \sigma^*(A_2) + 2k^2 \leq 2k^2 - 2k + 2k^2 = 4k^2 - 2k$$

as desired. If equality occurs, then $\sigma^*(A_2) = 2k^2 - 2k$ and A_2 has exactly $4k$ zeros, with k of them above the main diagonal. With the structure already determined for A , it is now straightforward to verify that A is a hollowed-diamond ASM of type 1 or type 2.

Case 2: $\sigma(Y) = 1$: In this case,

$$\sigma^*(A) \leq \sigma^*(A_2) + 2k^2 - 1 \leq 2k^2 - 2k + 1 + 2k^2 - 1 = 4k^2 - 2k$$

as desired. If equality occurs, then $\sigma^*(A_2) = 2k^2 - 2k + 1$ and A_2 has exactly $2(2k - 1)$ zeros, with $k - 1$ above the main diagonal. Again, with the structure already determined for A , it is straightforward to check that A is a hollowed-near-diamond ASM. \square

We now assume that $n \equiv 2 \pmod 4$. In this case, the diamond ASM D_n with $n = 4k + 2$ is a symmetric matrix whose main diagonal consists of $k + 1$ consecutive 0s, followed by $2k$ consecutive -1 s, followed by $k + 1$ consecutive 0s. Taking these $2k$ consecutive -1 s and every other pair of $+1$ s on the superdiagonal and subdiagonal, we get k disjoint principal 2×2 submatrices of the form

$$\begin{bmatrix} - & + \\ + & - \end{bmatrix}. \tag{21}$$

Replacing each these 2×2 submatrices of D_n with a 2×2 zero matrix, we get a symmetric ASM D_n^* with a zero main diagonal. If we take the rows of D_n in the reverse order, the -1 s on the main diagonal are replaced with $+1$ s, and the 2×2 submatrix (21) is also reversed. Continuing with the terminology used in the case $n = 4k$, we call both of these matrices *hollowed-diamond ASMs* of order $4k + 2$. For example, with $k = 2$ we have

This implies that the last row of X and the first column of Y each contain $2k$ nonzero entries. Thus, since each row of an ASM contains an odd number of nonzero entries, row $k + 1$ and column $3k + 2$ of A each contain the maximum number $2k + 1$ of nonzero entries. Hence for each i with $1 \leq i \leq k + 1$, row i contains the maximum number $2i - 1$ of nonzero entries and column $3k + 1 + i$ contains the maximum number $2i - 1$ nonzero entries. Thus the first $k + 1$ rows of A , after deletion of zero columns, form a $(k + 1) \times (2k + 1)$ matrix of the form in (20). This leads to a contradiction as in the proof of Theorem 5.1. Thus we must have $\sigma(Y) = 1$ (as in D_{4k+2}^*).

Now with $\sigma(Y) = 1$, as before we conclude that A has $2i - 1$ nonzeros in rows i and columns $4k + 3 - i$ for $i = 1, 2, \dots, k + 1$. Moreover, the $2i - 1$ nonzeros in these rows and columns must be above the main diagonal, as in D_{4k+2}^* . This allows us to conclude that $A = D_{4k+2}^*$ and A has $\frac{n^2 - 2n + 4}{2}$ nonzero entries where $n = 4k + 2$. The theorem now follows. \square

If we drop the symmetry assumption, it is not difficult to determine the maximum number of nonzero entries in an ASM whose main diagonal contains only zeros. In fact, this maximum is

$$\binom{n}{2} \text{ if } n \equiv 0, 3 \pmod{4}$$

and

$$\binom{n}{2} - 1 \text{ if } n \equiv 1, 2 \pmod{4}.$$

We illustrate ASMs that achieve these maxima for $n = 9, 10, 11$, and 12 from which the general pattern can be discovered.

$(n = 9)$

0				+				
	0		+	-	+			
		+	0	-	+	-	+	
	+	-	+	0	-	+	-	+
		+	-	+	0	-	+	
			0	+	-	+	0	
					+	-	+	0
						+		0

(35 nonzeros)

$(n = 10)$

0					+				
	0			+	-	+			
		0	+	-	+	-	+		
			0	0	+	-	+	-	+
		+	-	0	+	-	+	-	+
	+	-	+	-	0	+	-	+	
+	-	+	-	+	-	0	+		
	+	-	+	-	+	0	0		
			+	-	+				0
				+					0

(44 nonzeros)

ASMs; this can be argued by induction using the fact that an ASM with +1s in opposite corners has no other nonzeros in rows and columns 1 and n . It can also be checked that

$$\begin{bmatrix} & & & + \\ + & & & \\ & & & + \\ & & + & \\ + & & & \end{bmatrix}$$

is a maximal ASM. The permutation matrix

$$\begin{bmatrix} & + & & \\ & & & + \\ + & & & \\ & & & + \\ & & + & \end{bmatrix}$$

is not a maximal ASM, since

$$\begin{bmatrix} & + & & \\ & & & + \\ + & - & + & \\ & + & - & + \\ & & + & \end{bmatrix}$$

is an ASM obtained by replacing the 2×2 zero matrix in rows 3 and 4, and columns 2 and 3 with

$$\begin{bmatrix} - & + \\ + & - \end{bmatrix}. \tag{24}$$

Generalizing this example, let P be an $n \times n$ permutation matrix corresponding to the permutation π of $\{1, 2, \dots, n\}$. Let $1 \leq p < q \leq n$ and $1 \leq k < l \leq n$. Let $T_{p,q;k,l} = [t_{ij}]$ be the $n \times n$ $(0, +1, -1)$ -matrix such that

$$t_{ij} = \begin{cases} +1 & \text{if } (i, j) = (p, k) \text{ or } (q, l) \\ -1 & \text{if } (i, j) = (p, l) \text{ or } (q, k) \\ 0 & \text{otherwise.} \end{cases}$$

Suppose that the 2×2 submatrix determined by rows p and q and columns k and l of the ASM P is a zero matrix. Then $P + T_{p,q;k,l}$ is an ASM extension of P provided that there is a 1 of P above and to the right of the -1 in position (p, l) , and below and to the left of the -1 in position (q, k) , that is,

$$\pi(p) > l, \quad \pi(q) < k, \quad \pi^{-1}(k) > q, \quad \text{and} \quad \pi^{-1}(l) < p.$$

Pictorially, we have

	(k)	(l)	
		⊕	
(p)	0	0	⊕
(q) ⊕	0	0	
			⊕

where the 0s are in rows p and q and columns k and l , and the \oplus s indicate the relative positions of the $+$ s of P in rows p and q and columns k and l . Replacing $T_{p,q;k,l}$ with $-T_{p,q;k,l}$, we get a similar picture with the permutation π satisfying

$$\pi(p) < k, \quad \pi(q) > l, \quad \pi^{-1}(k) < p, \quad \text{and} \quad \pi^{-1}(l) > q.$$

Under these circumstances, we call $P \pm T_{p,q;k,l}$ an elementary ASM extension of P . Thus if B is an elementary ASM extension of P , then B has $n + 4$ nonzero entries of which exactly two are -1 .

We now show that if a permutation matrix P has an ASM extension, then it has an elementary ASM extension.

Theorem 6.1. *Let P be an $n \times n$ permutation matrix such that P is not a maximal ASM. Then P has an elementary ASM extension.*

Proof. Let B be an ASM extension of P where, as above, we distinguish each $+1$ of P by \oplus . Consider the $(0, +1, -1)$ -matrix $E = B - P$. We construct a bipartite graph G , the bipartition of whose vertices is given by the set U of positions in which E has $+1$ s and the set V of positions in which E has -1 s. There are an equal number of $+1$ s and -1 s in each row and in each column of E so that $|U| = |V|$. In each row and in each column of B a \oplus , if not the first or last $+$ in its row or column, lies between two consecutive -1 s of B , and thus on each of the four sides (horizontally and vertically) of such a \oplus , there are an equal number of $+1$ s and -1 s of B ; the $+1$ s and -1 s alternate from $+1$ to -1 to the left (respectively, above) a \oplus and from -1 to $+1$ to the right (respectively, below) a \oplus . The horizontal edges of G are obtained as follows: (i) by joining by an edge the position of a $+1$ which is to the left of the \oplus in its row to the position containing the -1 in its row that follows it, and (ii) by joining by an edge the position of a -1 which is to the right of the \oplus in its row to the position of the $+1$ that follows it. The vertical edges of G are similarly defined. Thus both the horizontal edges and the vertical edges form perfect matchings of G , and each vertex of G has degree equal to 2, meeting exactly one horizontal edge and exactly one vertical edge. It follows that the edges of G can be partitioned into cycles $\gamma_1, \dots, \gamma_k$, whose edges alternate between horizontal and vertical. This is illustrated in Figs. 2 and 3.

Continuing, let C_i be the $(0, +1, -1)$ -matrix whose nonzero entries are those corresponding to the vertices of γ_i ($i = 1, 2, \dots, k$). Then it follows that $P + C_i$ is an ASM which is an extension of P ; in fact, $P + \sum_{i \in K} C_i$ is an extension of P for all $\emptyset \neq K \subseteq \{1, 2, \dots, k\}$. Thus we may now assume that $B = P + C_1$, that is, G is a cycle of even length.

Consider the top row of C_1 that contains a position (p, q) with a -1 (so there is a \oplus above it) and assume without loss of generality that the edge of γ_1 goes to a $+1$ to its right. We follow γ_1 until we arrive at the first position (r, s) with a -1 for which the edge of γ_1 goes to the left; such a position exists since γ_1 is a cycle. There are two possibilities according to whether we arrive at this position (r, s) from above it or below it. See Fig. 2.

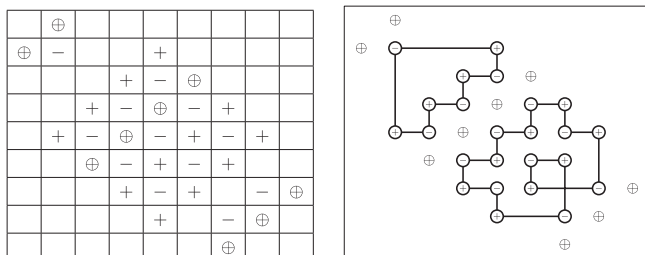


Fig. 2. An elementary ASM extension.

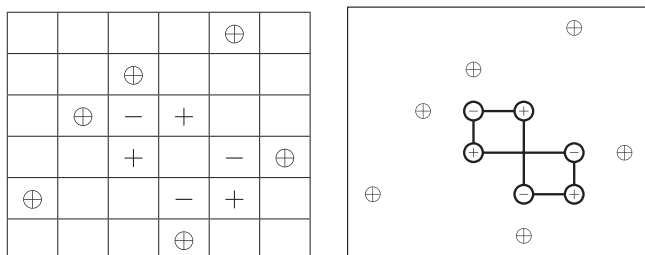


Fig. 3. An elementary ASM extension.

If we arrive from above, then the initial -1 in position (p, q) along with the -1 in the position (r, s) give the matrix $-T_{p,r;q,s}$. Considering the positions containing a \oplus , we see that $P - T_{p,r;q,s}$ is an elementary extension of P . Suppose we arrive at position (r, s) from below it. Let (g, h) be the position containing a -1 that came before (r, s) . Then considering again where the \oplus s are, we see that $P + T_{g,r;h,s}$ is an elementary extension of P . See Fig. 3. \square

Corollary 6.2. Let P be an $n \times n$ permutation matrix corresponding to the permutation π of $\{1, 2, \dots, n\}$. Then P is a maximal ASM if and only if there do not exist integers p, q, k, l with $1 \leq p < q \leq n$ and $1 \leq k < l \leq n$ such that

$$\pi(p) > l, \quad \pi(q) < k, \quad \pi^{-1}(k) > q, \quad \text{and} \quad \pi^{-1}(l) < p$$

or

$$\pi(p) < k, \quad \pi(q) > l, \quad \pi^{-1}(k) < p, \quad \text{and} \quad \pi^{-1}(l) > q.$$

Elementary extensions are obtained by appropriately adding $\pm T_{p,q;k,l}$ to an ASM A in such a way that the nonzero positions of A and those of $T_{p,q;k,l}$ do not overlap. It is also possible to add $\pm T_{p,q;k,l}$ to an ASM where the nonzero positions overlap and the result is an ASM. For example,

$$\begin{bmatrix} +1 & 0 & 0 \\ 0 & +1 & 0 \\ 0 & 0 & +1 \end{bmatrix} + \begin{bmatrix} -1 & +1 & 0 \\ +1 & -1 & 0 \\ 0 & 0 & 0 \end{bmatrix} = \begin{bmatrix} 0 & +1 & 0 \\ +1 & 0 & 0 \\ 0 & 0 & +1 \end{bmatrix}.$$

In fact, with two overlapping positions and starting with an ASM equal to the permutation matrix P as in the example, the result is another permutation matrix Q if and only if the permutation corresponding to Q is obtained from the permutation corresponding to P by a transposition. Another example, this time with an overlap of one position, is

$$\begin{bmatrix} 0 & 0 & +1 & 0 & 0 \\ 0 & +1 & 0 & 0 & 0 \\ +1 & 0 & -1 & 0 & +1 \\ 0 & 0 & 0 & +1 & 0 \\ 0 & 0 & +1 & 0 & 0 \end{bmatrix} + \begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & +1 & 0 \\ 0 & 0 & +1 & -1 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 0 & +1 & 0 & 0 \\ 0 & +1 & -1 & +1 & 0 \\ +1 & 0 & 0 & -1 & +1 \\ 0 & 0 & 0 & +1 & 0 \\ 0 & 0 & +1 & 0 & 0 \end{bmatrix}.$$

The above examples lead us to to define an *ASM-interchange* to be the operation (or result thereof) of adding $\pm T_{p,q;r,s}$ to an ASM provided the result is also an ASM. As shown above, ASM-interchanges generalize transpositions of permutations.

We now show that every ASM can be generated from the special ASM I_n by a sequence of ASM-interchanges. Since we can generate all permutation matrices from I_n by transpositions, this is equivalent to showing that every ASM results from the special subclass of ASMs consisting of permutation matrices by a sequence of ASM-interchanges. In fact, we can get every ASM from a permutation matrix by ASM-interchanges with overlap of only 1 or 2.

Theorem 6.3. *Let A be an $n \times n$ ASM. Then there is a sequence of ASM-interchanges such that starting from the identity matrix I_n we obtain A .*

Proof. If A does not have any -1 s, then A is a permutation matrix and we are done. Suppose $A = [a_{ij}]$ has at least one -1 , and choose the first row that has a -1 and the first column in that row that has a -1 . Suppose this -1 is in position (q, l) , that is, $a_{ql} = -1$. Since A is an ASM, there exists unique $p < q$ such that $a_{pl} = +1$ and $a_{il} = 0$ for all $i \neq p$ with $i < q$. Similarly, there exists unique $k < l$ such that $a_{qk} = +1$ and $a_{qj} = 0$ for all $j \neq k$ with $j < l$. Then $A + T_{p,q;k,l}$ is an ASM with one less -1 than A . Proceeding inductively, there is a sequence of ASM-interchanges which reduces A to an ASM B without any -1 s. Thus B is a permutation matrix, and further ASM-interchanges (i.e. transpositions) reduce A to I_n . \square

7. Concluding remarks

We conclude with some comments and open questions. As we have observed, every $(0, +1, -1)$ -matrix (possibly rectangular) is a submatrix of some ASM, and every symmetric $(0, 1, -1)$ -matrix is a principal submatrix of some symmetric ASM. If B is a $(0, +1, -1)$ -matrix, we define $\zeta(B)$ to be the smallest n such that B is a submatrix of an $n \times n$ ASM, and

$$\zeta(k, l) = \max\{\zeta(B) : B \text{ a } k \times l \text{ } (0, 1, -1)\text{-matrix}\} \quad (k, l \geq 1).$$

Note that if B is a submatrix of an $n \times n$ ASM, then by taking a direct sum with an identity matrix, B is a submatrix of an $m \times m$ ASM for all $m \geq n$.

Similarly, if C is a symmetric $(0, +1, -1)$ -matrix, we define $\zeta_p(C)$ to be the smallest n such that C is a principal submatrix of an $n \times n$ symmetric ASM, and

$$\zeta_p(k) = \max\{\zeta_p(C) : C \text{ a } k \times k \text{ symmetric } (0, 1, -1)\text{-matrix}\} \quad (k \geq 1).$$

Let

$$B = \begin{bmatrix} - & - \\ - & + \end{bmatrix}.$$

Then

$$A = \begin{bmatrix} & & & + & \\ & & & & \\ & + & & & \\ + & - & + & - & + \\ & & & & \\ & & & + & \\ + & & & & \end{bmatrix}$$

is an ASM whose submatrix determined by rows 3 and 4 and columns 2 and 4 equals B . Although B is not symmetric, we can insert a new row and column to get an ASM A' which contains B as a principal submatrix in rows and columns 3 and 5:

$$\begin{bmatrix} & & & & + & \\ & & & + & & \\ & & + & - & + & - & + \\ + & & & & & & \\ & & & & & + & \\ & & + & & & & \end{bmatrix}$$

Question 1. What are the values of $\zeta(k, l)$ and $\zeta_p(k)$, and which $(0, +1, -1)$ -matrices achieve these maxima?

Let J_k be the $k \times k$ matrix all of whose entries equal $+1$. Then J_k is a (principal) submatrix of the diamond ASM D_{4k-3} and $-J_k$ is a (principal) submatrix of the diamond ASM D_{4k-1} . It is natural to conjecture that $\zeta(k, k) = 4k - 1$, equivalently, that every $k \times k$ $(0, +1, -1)$ -matrix is a submatrix of a $(4k - 1) \times (4k - 1)$ ASM.

In Section 6 we have seen that a permutation matrix may or may not be maximal, that is, may or may not have an ASM extension.

Question 2. Characterize maximal ASMs.

Let A be an $n \times n$ ASM. As previously remarked, the seven non-identity row and column permutations of A corresponding to the dihedral group of order 8 result in ASMs, some of which may equal A . Let $\mathcal{C}(A)$ denote the equivalence class of all ASMs that can be gotten from A by permuting rows and columns. Thus if A is a permutation matrix, then $\mathcal{C}(A)$ is the set of all permutation matrices and has cardinality $n!$.

Question 3. What is

$$\max\{|\mathcal{C}(A)| : A \text{ an } n \times n \text{ ASM}\} \tag{25}$$

From the above, this maximum is at least $n!$ but, in general, it is much larger. For example, consider the $n \times n$ ASM A which is the direct sum of l copies of D_3 and one I_1 (so $n = 3l + 1$). By recursively choosing three rows and three columns for each of the $l D_3$ s to go into an $n \times n$ matrix and then noting that each set of choices can be gotten in $l!$ ways, we see that the number of ASMs in $\mathcal{C}(A)$ is

$$\frac{1}{l!} \binom{n}{3}^2 \binom{n-3}{3}^2 \cdots \binom{4}{3}^2 = \frac{n!^2}{l!6^{2l}} \tag{26}$$

If $l \geq 2$, then (26) is greater than $n!$. If A' is the direct sum of $l - 1$ copies of D_3 and one D_4 , the number of ASMs in $\mathcal{C}(A')$ is

$$2 \frac{1}{(l-1)!} \binom{n}{3}^2 \binom{n-3}{3}^2 \cdots \binom{7}{3}^2 = \frac{2n!^2}{(l-1)!6^{2l}16}. \quad (27)$$

(The factor of 2 is due to the fact that the rows of D_4 can be reversed to create the other matrix we have referred to as D_4 .) Subtracting (26) from (27), we see that if $l > 8$,

$$|\mathcal{C}(A')| > |\mathcal{C}(A)|.$$

These calculations illustrate that the determination of (25) is apt to be very difficult. A related question is:

Question 4. What is

$$\max\{|\mathcal{C}(A)| : A \text{ an } n \times n \text{ connected ASM}\}?$$

Acknowledgements

We are indebted to two anonymous referees who read the paper very carefully and made many suggestions for improving the clarity of our paper.

References

- [1] D. Bressoud, Proofs and Confirmations: The Story of the Alternating Sign Conjecture, Math. Association of America, Cambridge University Press, Cambridge, 1999.
- [2] R.A. Brualdi, H.J. Ryser, Combinatorial Matrix Theory, Cambridge University Press, Cambridge, 1991.
- [3] G. Kuperberg, Another proof of the alternating sign matrix conjecture, Inter. Math. Res. Notes 1996 (1996) 139–150.
- [4] W.H. Mills, D.P. Robbins, H. Rumsey, Alternating-sign matrices and descending plane partitions, J. Combin. Theory Ser. A 34 (1983) 340–359.
- [5] J. Propp, The many faces of alternating sign matrices, Discrete models: combinatorics, computation, and geometry (Paris, 2001) 043058 (electronic), Discrete Math. Theor. Comput. Sci. Proc., AA, Maison Inform. Math. Discrèt. (MIMD), Paris, 2001.
- [6] D.P. Robbins, The story of 1, 2, 7, 42, 429, 7436, ..., Math. Intelligencer 13 (1991) 12–19.
- [7] D. Zeilberger, Proof of the alternating sign matrix conjecture, Electron. J. Combin. 3 (1996) R13